

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ МИКРО-ВЫРАЖЕНИЙ ЛИЦА

*В.Г. Спицын, д.т.н., проф. ОИТ ИШИТР,
В.А. Коровкин, аспирант гр. А7-39,
В.В. Видман, аспирант гр. А7-39
Томский Политехнический университет
E-mail: alcasar@tpu.com*

Введение

Выражения лица человека можно разделить на два типа «макро-выражения» и «микро-выражения», которые показывают эмоциональное состояние человека в конкретной момент времени. Макро-выражения — это обычные выражения, которые встречаются при повседневном взаимодействии с людьми и длятся от 0,5 до 4 секунд. В большинстве случаев они легко считываются без подготовки визуальным способом.

В критических или в опасных (волнительных, конфликтных и т.д.) ситуациях люди часто пытаются скрыть или подавить свои истинные эмоции [1]. Эти скрытые эмоции рефлекторно проявляются в течение от 0,2 до 0,4 секунд и называются «микро-выражениями». Микро-выражения были первоначально обнаружены Хаггардом и Айзексом [2], а затем подтверждены в работе Экманом и Фризеном [3]. Помимо короткой продолжительности, они также обладают низкой интенсивностью. Эти две уникальные особенности делает распознавание микро-выражений людей сложной задачей даже для хорошо обученных специалистов.

Ограниченная способность людей эффективно распознавать микро-выражения требует разработки специализированных алгоритмов, которые смогут их автоматически распознавать. Микровыражения могут показывать такие эмоции как печаль, счастье, страх, гнев, удивление, отвращение и презрение.

Степень разработанности темы

Сегодня можно найти много исследований и алгоритмов для детектирования и классификации макро-выражений на основе эмоций, однако в области определения микро-выражений было проведено лишь несколько исследований. В существующих исследованиях использовались алгоритмы контролируемого машинного обучения, которые включают деревья решений, нейронные сети, метод k-ближайших соседей (KNN), машину опорных векторов (SVM), extreme learning method (ELM) и классификаторы случайных леса (random forest classifier). SVM также широко используется работы с микро-выражениями, как и для других задач классификации, в то время как ELM не используется многими для тех же целей. Машина опорных векторов, по-видимому, является наиболее часто используемой моделью машинного обучения для распознавания выражений лица из-за ее хорошей предсказуемой производительности, обобщения независимо от систематической ошибки в обучающей выборке. Однако у SVM есть несколько серьезных недостатков, в том числе их сложность и низкая скорость обучения.

Метод опорных векторов в работе используется в качестве базовой модели распознавания для сравнения. Главная цель проводимых экспериментов решить проблему низкой скорости его обучения.

Для извлечения признаков для микровыражений, то использовался метод локальных бинарных шаблонов (LBP), а именно метода Local Binary Pattern histograms from Three Orthogonal Planes (LBP-TOР).

Производительность двух моделей (SVM и ELM) сравнивается как по статическим, так и по временным характеристикам, и сравнивается их общее время обучения. Модели оценивались с использованием специального датасета микроэкспрессий CASME II (видео) и CK+(изображения). Кроме нее существуют и другие базы данных – USFHD или SMIC. На текущий момент китайская CASME II является последней, актуальной и общедоступной для исследований.

Датасет CASME II

Датасет состоит из 247 образцов микро-выражений лица, полученных от 26 участников, которые смотрели эмоциональные видеоклипы. Во время просмотра видео перед каждым из участников помещался экран, а для записи их эмоций использовалась камера с высоким разрешением. После просмотра клипов участникам было предложено оценить эмоциональную интенсивность видеоклипов по 7-балльной шкале Ликерта, где 0 – самый низкий, а 6 – самый высокий. В свою очередь видеозаписи, полученные от каждого участника, были разделены по карам и обработаны (удаление нерегулярных

движений лица и тела). В результате были получены образцы микро-выражений, которые имели общую продолжительность менее 500 миллисекунд. Эти образцы были обработаны экспертами с помощью метода СКЛиД. Однако стоит иметь ввиду, что критерии маркировки микро-выражений полностью не совпадают с критериями обычных эмоций. Образцы микро-выражения состоят из пяти классов, которые включают отвращение, счастье, удивление, печаль и другие. Точное детектирование остальных не представлялось возможным. Таким образом, датасет состоял из 247 образцов: счастьеб отвращение, сюрприз, печаль и другие. Каждая видео последовательность имеет начальный кадр и Арех-кадр.

Начальный кадр — это первый кадр, в котором происходит изменение нейтрального выражения.

Кадр Арех — это кадр, в котором достигается наивысшая интенсивность выражения, в то время как кадр смещения является последним кадром перед изменением выражения лица на нейтральное.

В итоге, CASME II содержит в общей сложности 247 образцов микровыражений с частотой дискретизации 200 кадров в секунду.

Извлечение признаков объектов из последовательности изображений (временные данные) осуществлялась с помощью алгоритма LBP-TOP. Он является одним из пространственно-временных дескрипторов для динамических текстур (то есть текстур в движении), который был создан для того, чтобы преодолеть недостатки обычного LBP. Основным недостатком LBP является то, что он может извлекать функции только из неподвижных изображений. LBP-TOP является одним из вариантов обычных LBPS и был предложен Чжао и Пиетикайненом для анализа текстур, зависящих от времени (или анализа видео).

Эксперимент

Обучение SVM на арех-кадрах проводилось путем загрузки (1x177) вектора признаков для 220 образцов, полученных после извлечения признаков LBP. После этого была проведена пятикратная перекрестная проверка, чтобы разделить образцы на пять независимых подмножеств. Для обучения использовалось линейное ядро SVM. Средняя точность обучения 96,30% была достигнута с помощью функций LBP.

Таблица 1. Результаты обучения SVM

| Класс микро-выражения | Точность при LBP | Точность при LBP-TOP |
|-----------------------|------------------|----------------------|
| Отвращение | 94,99 | 91,74 |
| Счастье | 97,21 | 87,83 |
| Подавленность | 96,89 | 90,44 |
| Удивление | 98,01 | 88,70 |
| Другие | 93,01 | 91,30 |
| Средние | 96,3 | 91,74 |

Финальная архитектура ELM представлена в таблице ниже:

Таблица 2. Архитектура ELM

| Детали | Арех кадры | Последовательность изображений |
|-------------------------|----------------|--------------------------------|
| Количество примеров | 220 | 230 |
| Feature Vector Size | 1 × 256 | 1 × 177 |
| ELM Model Type | Classification | Classification |
| Число входных нейронов | 256 | 177 |
| Число выходных нейронов | 2 | 2 |
| Label | 1 (Позитивный) | 1 (позитивный) |
| | 0 (негативный) | 0 (негативный) |

В ходе проведения ряда численных экспериментов были получены следующие усредненные данные:

Таблица 3 Сравнение методов машинного обучения в задаче классификации микро-выражений

| Метод машинного обучения | Распознаваемая эмоция | | | | | |
|--------------------------|-----------------------|---------|-----------|-----------|--------|---------|
| | отвращение | счастье | репрессия | удивление | другие | Среднее |
| SVM | 94.99 | 85.31 | 88.45 | 92.24 | 87.27 | 89.65 |
| ELM | 95.22 | 86.96 | 87.83 | 93.74 | 85.59 | 89.87 |

Таблица 4. Сравнение методов машинного обучения в задаче классификации микро-выражений повремени обучения (в секундах)

| Класс микро-выражения | SVM | ELM |
|------------------------|--------|--------|
| Отвращение | 0.2808 | 0.0468 |
| Радость | 0.3806 | 0.0499 |
| Грусть | 0.2434 | 0.0593 |
| Удивления | 0.2995 | 0.0530 |
| Другие | 0.3182 | 0.0406 |
| Среднее время обучения | 0.3405 | 0.0499 |

Из результатов эксперимента видно, что метод ELM значительно превосходит по скорости обучения традиционный метод SVM. При этом сохраняется сравнительно высокая точность работы.

Заключение

В результате проведенной работы был разработан метод машинного обучения на основе машины экстремального обучения (EML) и бинарных локальных шаблонов (LBP и TOP-LBP). Эксперименты показали высокую скорость обучения предложенной архитектуры по сравнению с часто применяемым SVM при приемлемой оценке точности. Метод себя показал одинаково хорошо, как и на статичных изображениях, так и на видеоряде.

Список литературы

1. L. Su, M.D. Levine. High-stakes deception detection based on facial expressions // 22nd IEEE International Conference on Pattern Recognition (ICPR). – 2014. – P. 2519-2524.
2. E.A. Haggard, K.S. Isaacs. Micro-momentary facial expressions as indicators of ego-mechanisms in psychotherapy // Methods of Research in Psychotherapy. – Springer. –1966. –P. 154-165
3. P. Ekman, W.V. Friesen. Non-verbal leakage and clues to deception. // Psychiatry. – 1969. – V. 32.– №1. – P. 88-106
4. Коровкин В. А. Применение нейронных сетей для классификации эмоции на изображениях // XX всероссийская конференция молодых учёных по математическому моделированию и информационным технологиям УМ-2019, Новосибирск, 28 октября - 1 ноября 2019. – Новосибирск: ИВТ СО РАН. — С. 62-63
5. Коровкин В.А. Применение методов машинного обучения для решения задачи классификации эмоции на изображении по ключевым точкам // Информационные технологии в науке, управлении, социальной сфере и медицине. Сборник научных трудов VI Международной научной конференции. Под редакцией О.Г. Берестневой, В.В. Спицына, А.И. Труфанов, Т.А. Гладковой. 2019. С. 100-104
6. Коровкин В. А. Распознавание и классификация лицевых эмоций на основе визуальной информации на видеопотоке // Молодежь и современные информационные технологии: сборник трудов XVII Международной научно-практической конференции студентов, аспирантов и молодых ученых (Томск, 17–20 февраля 2020 г.) / Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2020. – С. 39-40